

```
Bi
```

```
gatccaatgt ttaaatatca cctcactgtg gcccaggttc gaggagggat 1350
ggtgtttgag ctagccaatt ccatagtgct cccttttgat tgtcgagatt 1400
atgctgtagt tttaagaaag tatgctgaca aaatctacaa tatttctatg 1450
aaacatccac aggaaatgaa gacatacagt ttatcatttg attcactttt 1500
ttctgcagta aaaaatttta cagaaattgc ttccaagttc agcgagagac 1550
tccaggactt tgacaaaagc aacccaatat tgttaagaat gatgaatgat 1600
caactcatgt ttctggaaag agcatttatt qatccattag ggttaccaga 1650
cagacetttt tataggeatg teatetatge tecaageage caeaacaagt 1700
atgcagggga gtcattccca ggaatttatg atgctctgtt tgatattgaa 1750
agcaaagtgg accettecaa ggcetgggga gatgtgaaga gacagattte 1800
tgttgcagcc ttcacagtgc aggcagctgc agagactttg agtgaagtag 1850
cctaagagga ttctttagag actctgtatt gaatttgtgt ggtatgtcac 1900
tcaaagaata ataatgggta tattgataaa ttttaaaatt ggtatatttg 1950
1992
     <210>
     <211>
               442
     <212>
               PRT
     <213>
               Homo sapiens
     <220>
     <223>
               deduced amino acid sequence of PSMA-like
               protein
     <400>
Met Gly Gly Ser Ala Pro Pro Asp Ser Ser Trp Arg Gly Ser Leu
                5
                                    10
                                                        15
Lys Val Ser Tyr Asn Val Gly Pro Gly Phe Thr Gly Asn Phe Ser
                                                        30
                20
                                    25
Thr Gln Lys Val Lys Met His Ile His Ser Thr Asn Glu Val Thr
                                    40
                                                        45
Arg Ile Tyr Asn Val Ile Gly Thr Leu Arg Gly Ala Val Glu Pro
                50
                                    55
                                                        60
Asp Arg Tyr Val Ile Leu Gly Gly His Arg Asp Ser Trp Val Phe
                65
                                    70
                                                        75
Gly Gly Ile Asp Pro Gln Ser Gly Ala Ala Val Val His Glu Thr
                                    85
                                                        90
                80
Val Arg Ser Phe Gly Thr Leu Lys Lys Glu Gly Trp Arg Pro Arg
                95
                                    100
                                                        105
Arg Thr Ile Leu Phe Ala Ser Trp Asp Ala Glu Glu Phe Gly Leu
                                    115
                                                        120
                110
Leu Gly Ser Thr Glu Trp Ala Glu Asp Asn Ser Arg Leu Leu Gln
                125
                                    130
                                                        135
```

BI

Glu Arg Gly Val Ala Tyr Ile Asn Ala Asp Ser Ser Ile Glu Gly Asn Tyr Thr Leu Arg Val Asp Cys Thr Pro Leu Met Tyr Ser Leu Val Tyr Asn Leu Thr Lys Glu Leu Lys Ser Pro Asp Glu Gly Phe Glu Gly Lys Ser Leu Tyr Glu Ser Trp Thr Lys Lys Ser Pro Ser Pro Glu Phe Ser Gly Met Pro Arg Ile Ser Lys Leu Gly Ser Gly Asn Asp Phe Glu Val Phe Phe Gln Arg Leu Gly Ile Ala Ser Gly Arg Ala Arg Tyr Thr Lys Asn Trp Glu Thr Asn Lys Phe Ser Gly Tyr Pro Leu Tyr His Ser Val Tyr Glu Thr Tyr Glu Leu Val Glu Lys Phe Tyr Asp Pro Met Phe Lys Tyr His Leu Thr Val Ala Gln Val Arg Gly Gly Met Val Phe Glu Leu Ala Asn Ser Ile Val Leu Pro Phe Asp Cys Arg Asp Tyr Ala Val Val Leu Arg Lys Tyr Ala Asp Lys Ile Tyr Asn Ile Ser Met Lys His Pro Gln Glu Met Lys Thr Tyr Ser Leu Ser Phe Asp Ser Leu Phe Ser Ala Val Lys Asn Phe Thr Glu Ile Ala Ser Lys Phe Ser Glu Arg Leu Gln Asp Phe Asp Lys Ser Asn Pro Ile Leu Leu Arg Met Met Asn Asp Gln Leu Met Phe Leu Glu Arg Ala Phe Ile Asp Pro Leu Gly Leu Pro Asp Arg Pro Phe Tyr Arg His Val Ile Tyr Ala Pro Ser Ser His Asn

```
Lys Tyr Ala Gly Glu Ser Phe Pro Gly Ile Tyr Asp Ala Leu Phe
                395
                                    400
                                                         405
Asp Ile Glu Ser Lys Val Asp Pro Ser Lys Ala Trp Gly Asp Val
                410
                                    415
                                                         420
Lys Arg Gln Ile Ser Val Ala Ala Phe Thr Val Gln Ala Ala Ala
                425
                                    430
                                                         435
Glu Thr Leu Ser Glu Val Ala
                440
     <210>
     <211>
               2653
     <212>
               DNA
     <213>
               Homo sapiens
     <220>
               nucleotide sequence of human PSMA gene
     <223>
     <300>
               GenBank Accession No. M99487
     <308>
     <309>
     <400>
               3
ctcaaaaggg gccggatttc cttctcctgg aggcagatgt tgcctctctc
                                                          50
tctcgctcgg attggttcag tgcactctag aaacactgct gtggtggaga
                                                         100
aactggaccc caggtctgga gcgaattcca gcctgcaggg ctgataagcg
                                                         150
                                                         200
aggcattagt gagattgaga gagactttac cccgccgtgg tggttggagg
gcgcgcagta gagcagcagc acaggcgcgg gtcccgggag gccggctctg
                                                         250
ctcgcgccga gatgtggaat ctccttcacg aaaccgactc ggctgtggcc
                                                         300
accgcgcgcc gcccgcgctg gctgtgcgct ggggcgctgg tgctggcggg
                                                         350
tggcttcttt ctcctcggct tcctcttcgg gtggtttata aaatcctcca
                                                         400
atgaagctac taacattact ccaaagcata atatgaaagc atttttggat
                                                         450
gaattgaaag ctgagaacat caagaagttc ttatataatt ttacacagat
                                                         500
accacattta gcaggaacag aacaaaactt tcagcttgca aagcaaattc
                                                         550
aatcccagtg gaaagaattt ggcctggatt ctgttgagct agcacattat
                                                         600
gatgtcctgt tgtcctaccc aaataagact catcccaact acatctcaat
                                                         650
aattaatgaa gatggaaatg agattttcaa cacatcatta tttgaaccac
                                                         700
ctcctccagg atatgaaaat gtttcggata ttgtaccacc tttcagtgct
                                                         750
tteteteete aaggaatgee agagggegat etagtgtatg ttaaetatge
                                                         800
acgaactgaa gacttcttta aattggaacg ggacatgaaa atcaattgct
                                                         850
ctgggaaaat tgtaattgcc agatatggga aagttttcag aggaaataag
                                                         900
gttaaaaatg cccagctggc aggggccaaa ggagtcattc tctactccga
ccctgctgac tactttgctc ctggggtgaa gtcctatcca gatggttgga 1000
atcttcctgg aggtggtgtc cagcgtggaa atatcctaaa tctgaatggt 1050
gcaggagacc ctctcacacc aggttaccca gcaaatgaat atgcttatag 1100
gcgtggaatt gcagaggctg ttggtcttcc aagtattcct gttcatccaa 1150
ttggatacta tgatgcacag aagctcctag aaaaaatggg tggctcagca 1200
ccaccagata gcagctggag aggaagtctc aaagtgccct acaatgttgg 1250
```

```
acctggcttt actggaaact tttctacaca aaaagtcaag atgcacatcc 1300
actctaccaa tgaagtgaca agaatttaca atgtgatagg tactctcaga 1350
ggagcagtgg aaccagacag atatgtcatt ctgggaggtc accgggactc 1400
atgggtgttt ggtggtattg accetcagag tggagcaget gttgttcatg 1450
aaattgtgag gagctttgga acactgaaaa aggaagggtg gagacctaga 1500
agaacaattt tgtttgcaag ctgggatgca gaagaatttg gtcttcttgg 1550
ttctactgag tgggcagagg agaattcaag actccttcaa gagcgtggcg 1600
tggcttatat taatgctgac tcatctatag aaggaaacta cactctgaga 1650
gttgattgta caccgctgat gtacagcttg gtacacaacc taacaaaaga 1700
gctgaaaagc cctgatgaag gctttgaagg caaatctctt tatgaaagtt 1750
ggactaaaaa aagtccttcc ccagagttca gtggcatgcc caggataagc 1800
aaattgggat ctggaaatga ttttgaggtg ttcttccaac gacttggaat 1850
tgcttcaggc agagcacggt atactaaaaa ttgggaaaca aacaaattca 1900
gcggctatcc actgtatcac agtgtctatg aaacatatga gttggtggaa 1950
aagttttatg atccaatgtt taaatatcac ctcactgtgg cccaggttcg 2000
aggagggatg gtgtttgagc tagccaattc catagtgctc ccttttgatt 2050
gtcgagatta tgctgtagtt ttaagaaagt atgctgacaa aatctacagt 2100
atttctatga aacatccaca ggaaatgaag acatacagtg tatcatttga 2150
ttcacttttt tctgcagtaa agaattttac agaaattgct tccaagttca 2200
gtgagagact ccaggacttt gacaaaagca acccaatagt attaagaatg 2250
atgaatgatc aactcatgtt tctggaaaga gcatttattg atccattagg 2300
gttaccagac aggccttttt ataggcatgt catctatgct ccaagcagcc 2350
acaacaagta tgcaggggag tcattcccag gaatttatga tgctctgttt 2400
gatattgaaa gcaaagtgga cccttccaag gcctggggag aagtgaagag 2450
acagatttat gttgcagcct tcacagtgca ggcagctgca gagactttga 2500
gtgaagtagc ctaagaggat tctttagaga atccgtattg aatttgtgtg 2550
gtatgtcact cagaaagaat cgtaatgggt atattgataa attttaaaat 2600
2653
     <210>
     <211>
               750
     <212>
               PRT
     <213>
               Homo sapiens
     <220>
     <223>
               deduced amino acid sequence of PSMA protein
     <400>
Met Trp Asn Leu Leu His Glu Thr Asp Ser Ala Val Ala Thr Ala
```

B

Phe Leu Asp Glu Leu Lys Ala Glu Asn Ile Lys Lys Phe Leu Tyr Asn Phe Thr Gln Ile Pro His Leu Ala Gly Thr Glu Gln Asn Phe Gln Leu Ala Lys Gln Ile Gln Ser Gln Trp Lys Glu Phe Gly Leu Asp Ser Val Glu Leu Ala His Tyr Asp Val Leu Leu Ser Tyr Pro Asn Lys Thr His Pro Asn Tyr Ile Ser Ile Ile Asn Glu Asp Gly Asn Glu Ile Phe Asn Thr Ser Leu Phe Glu Pro Pro Pro Gly Tyr Glu Asn Val Ser Asp Ile Val Pro Pro Phe Ser Ala Phe Ser Pro Gln Gly Met Pro Glu Gly Asp Leu Val Tyr Val Asn Tyr Ala Arg Thr Glu Asp Phe Phe Lys Leu Glu Arg Asp Met Lys Ile Asn Cys Ser Gly Lys Ile Val Ile Ala Arg Tyr Gly Lys Val Phe Arg Gly Asn Lys Val Lys Asn Ala Gln Leu Ala Gly Ala Lys Gly Val Ile Leu Tyr Ser Asp Pro Ala Asp Tyr Phe Ala Pro Gly Val Lys Ser Tyr Pro Asp Gly Trp Asn Leu Pro Gly Gly Gly Val Gln Arg Gly Asn Ile Leu Asn Leu Asn Gly Ala Gly Asp Pro Leu Thr Pro Gly Tyr Pro Ala Asn Glu Tyr Ala Tyr Arg Arg Gly Ile Ala Glu Ala Val Gly Leu Pro Ser Ile Pro Val His Pro Ile Gly Tyr Tyr Asp Ala Gln Lys Leu Glu Lys Met Gly Gly Ser Ala Pro Pro



Asp Ser Ser Trp Arg Gly Ser Leu Lys Val Pro Tyr Asn Val Gly Pro Gly Phe Thr Gly Asn Phe Ser Thr Gln Lys Val Lys Met His Ile His Ser Thr Asn Glu Val Thr Arg Ile Tyr Asn Val Ile Gly Thr Leu Arg Gly Ala Val Glu Pro Asp Arg Tyr Val Ile Leu Gly Gly His Arg Asp Ser Trp Val Phe Gly Gly Ile Asp Pro Gln Ser Gly Ala Ala Val Val His Glu Ile Val Arg Ser Phe Gly Thr Leu Lys Lys Glu Gly Trp Arg Pro Arg Arg Thr Ile Leu Phe Ala Ser Trp Asp Ala Glu Glu Phe Gly Leu Leu Gly Ser Thr Glu Trp Ala Glu Glu Asn Ser Arg Leu Leu Gln Glu Arg Gly Val Ala Tyr Ile Asn Ala Asp Ser Ser Ile Glu Gly Asn Tyr Thr Leu Arg Val Asp Cys Thr Pro Leu Met Tyr Ser Leu Val His Asn Leu Thr Lys Glu Leu Lys Ser Pro Asp Glu Gly Phe Glu Gly Lys Ser Leu Tyr Glu Ser Trp Thr Lys Lys Ser Pro Ser Pro Glu Phe Ser Gly Met Pro Arg Ile Ser Lys Leu Gly Ser Gly Asn Asp Phe Glu Val Phe Phe Gln Arg Leu Gly Ile Ala Ser Gly Arg Ala Arg Tyr Thr Lys Asn Trp Glu Thr Asn Lys Phe Ser Gly Tyr Pro Leu Tyr His Ser Val Tyr Glu Thr Tyr Glu Leu Val Glu Lys Phe Tyr Asp Pro Met Phe

Lys	Tyr	His	Leu	Thr 575	Val	Ala	Gln	Val	Arg 580	Gly	Gly	Met	Val	Phe 585
Glu	Leu	Ala	Asn	Ser 590	Ile	Val	Leu	Pro		Asp	Cys	Arg	Asp	
Ala	Val	Val	Leu	Arg	Lys	Tyr	Ala	Asp		Ile	Tyr	Ser	Ile	
Met	Lys	His	Pro	Gln 620	Glu	Met	Lys	Thr	Tyr 625	Ser	Val	Ser	Phe	
Ser	Leu	Phe	Ser	Ala 635	Val	Lys	Asn	Phe	Thr 640	Glu	Ile	Ala	Ser	Lys 645
Phe	Ser	Glu	Arg	Leu 650	Gln	Asp	Phe	Asp	Lys 655	Ser	Asn	Pro	Ile	Val 660
Leu	Arg	Met	Met	Asn 665	Asp	Gln	Leu	Met	Phe 670	Leu	Glu	Arg	Ala	Phe 675
Ile	Asp	Pro	Leu	Gly 680	Leu	Pro	Asp	Arg	Pro 685	Phe	Tyr	Arg	His	Val 690
Ile	Tyr	Ala	Pro	Ser 695	Ser	His	Asn	Lys	Tyr 700	Ala	Gly	Glu	Ser	Phe
Pro	Gly	Ile	Tyr	Asp 710	Ala	Leu	Phe	Asp	Ile 715	Glu	Ser	Lys	Val	Asp 720
Pro	Ser	Lys	Ala	Trp 725	Gly	Glu	Val	Lys	Arg 730	Gln	Ile	Tyr	Val	Ala 735
Ala	Phe	Thr	Val	Gln 740	Ala	Ala	Ala	Glu	Thr 745	Leu	Ser	Glu	Val	Ala 750
	.0.	1.0.		_										
		10> 11>		5 28										
		12> 13>		DNA Arti	fici	al s	eque	nce						
	<2	20>												
		21>		prim			~ ~		a =-	.]:	mm 7 -	£	
	<2.	23>		sens the chro	firs	t in	tron							

28

<400> 5

gccttcattt tcagaacatc tcatgcat

	<210> <211> <212> <213>	
	<220> <221> <223>	primer_bind antisense primer designed for only amplifying the first intron of the PSMA- like gene on chromosome 11q
	<400> gtccatataa	6 actttcaaga atgtg 25
) ,	<210> <211> <212> <213>	
,	<220> <221> <223>	primer_bind sense oligonucleotide primer based upon intronic sequences of the PSMA genomic clone used to amplify the corresponding regions of the PSMA-like gene (exon 2)
	<400> ctcacctaat	7 gtcagaggta 20
	<210> <211> <212> <213>	
4	<220> <221> <223>	primer_bind antisense oligonucleotide primer based upor intronic sequences of the PSMA genomic clone used to amplify the corresponding regions of the PSMA-like gene (exon 2)
	<400> agtatagtcc	8 tcctcagatg 20
	<210><211><212><213>	9 24 DNA Artificial sequence

	<220> <221> <223>	primer_bind sense oligonucleotide primer based upon intronic sequences of the PSMA genomic clone used to amplify the corresponding regions of the PSMA-like gene (exon 3)
	<400> caaagtactt	9 ttgtgtaact ctgc 24
	<210> <211> <212> <213>	10 22 DNA Artificial sequence
1	<220> <221> <223>	primer_bind antisense oligonucleotide primer based upon intronic sequences of the PSMA genomic clone used to amplify the corresponding regions of the PSMA-like gene (exon 3)
•	<400> cataggaaag	10 tagttgacac gg 22
	<210> <211> <212> <213>	11 22 DNA Artificial sequence
••	<220> <221> <223>	primer_bind sense oligonucleotide primer based upon intronic sequences of the PSMA genomic clone used to amplify the corresponding regions of the PSMA-like gene (exon 4)
	<400> cctgaaggat	11 tcattcaccc tc 22
	<210> <211> <212> <213>	12 24 DNA Artificial sequence
	<220> <221> <223>	<pre>primer_bind antisense oligonucleotide primer based upon</pre>

intronic sequences of the PSMA genomic clone used to amplify the corresponding regions of the PSMA-like gene (exon 4)

		g (2.1211 1,
<400>	12	
	ttatcggctg aaca	24
gacccccaa	ccaceggeeg daed	24
<210>	13	
<211>		
<212>		
<213>	Arcificial sequence	
<220>		
<221>	primer_bind	
<223>	<u>-</u>	nrimer based upon
\4257	intronic sequences of	
	clone used to amplify	
	regions of the PSMA-li	
	regions of the PSMA-11	ke gene (exons 5-6)
<400>	13	
	gtccccatgc ag	22
acycecaaca	geceeatge ag	22
<210>	14	
<211>		
<212>		
<213>		
<213 <i>></i>	Artificial sequence	
<220>		
<221>	primer_bind	
<223>		ide primer based upon
\2237	intronic sequences of	
	clone used to amplify regions of the PSMA-li	
	regions of the PSMA-11	ke gene (exons 5-6)
<400>	14	
	gtccattgta cc	22
gacacgeeca	geceaeegea ee	22
<210>	15	
<211>	22	
<212>		
<213>	Artificial sequence	
<220>		
<221>	primer_bind	
<223>	sense oligonucleotide	primar based upon
<443>		
	intronic sequences of	<u> </u>
	clone used to amplify	
	regions of the PSMA-li	ke gene (exon /)

<400> gaaccgtttg	15 aatgaaactg ag	22
<210> <211> <212> <213>	22 DNA	
<220> <221> <223>	antisense oligonucl intronic sequences	fy the corresponding
<400> ttacccaaat	16 agccatccat gg	22
<210> <211> <212> <213>	23 DNA	
<220> <221> <223>	sense oligonucleotic intronic sequences clone used to ampli	de primer based upon of the PSMA genomic fy the corresponding -like gene (exons 8-9)
<400> gcagatgctc	17 aataagtgaa tcc	23
<210> <211> <212> <213>	18 24 DNA Artificial sequence	
<220> <221> <223>	antisense oligonucl intronic sequences clone used to ampli	eotide primer based upon of the PSMA genomic fy the corresponding -like gene (exons 8-9)
<400> ccagcacata	18 acagttactt gatc	24
<210>	19	

<211> <212> <213>	
<220> <221> <223>	-
<400>	19
tagatgctat	tgagtcgttt gc 22
<210> <211> <212> <213>	22 DNA
<220> <221> <223>	<u>. — </u>
<400>	20
aaactgagac	tcagataggc tg 22
<210><211><212><213>	DNA
<220> <221> <223>	primer_bind sense oligonucleotide primer based upon intronic sequences of the PSMA genomic clone used to amplify the corresponding regions of the PSMA-like gene (exon 11)
<400>	21 tagtgtcctg gg 22
	g-ggg
<210> <211> <212> <213>	22 24 DNA Artificial sequence

<220> <221> <223>	antisense oligonuci intronic sequences clone used to ampl	leotide primer based upor of the PSMA genomic ify the corresponding A-like gene (exon 11)
<400> gcttggcaaa	22 caagtcctgg ctac	24
<210> <211> <212> <213>	22 DNA	e
<220> <221> <223>	<pre>primer_bind sense oligonucleot: intronic sequences clone used to ampl:</pre>	ide primer based upon of the PSMA genomic ify the corresponding A-like gene (exon 12)
<400> tgtcgttaat	23 atgggtcagc tc	22
<210> <211> <212> <213>	22 DNA	e
<220> <221> <223>	<pre>primer_bind antisense oligonuc intronic sequences clone used to ampl</pre>	leotide primer based upor of the PSMA genomic ify the corresponding A-like gene (exon 12)
<400> ttaactagac	24 tgctgctcct ag	22
<210> <211> <212> <213>	22 DNA	e
<220> <221> <223>	<pre>primer_bind sense oligonucleot.</pre>	ide primer based upon of the PSMA genomic

clone used to amplify the corresponding regions of the PSMA-like gene (exon 13)

	J	, (See 1
<400> tggtaggaat	25 ttagcagtgg tc	22
<210> <211> <212> <213>	22 DNA	
<220> <221> <223>	primer_bind	y the corresponding
<400>		00
gatgctacta	atgggctacc tc	22
<210> <211> <212> <213>	22 DNA	
<220> <221> <223>	primer_bind	of the PSMA genomic Ty the corresponding
<400>	27	
cttctggtta	atggacatct ag	22
<210> <211> <212> <213>	DNA	
<220> <221> <223>	primer_bind	y the corresponding
<400> caatcccaca	28 ctgaattcag tg	22

<210> <211> <212> <213>	22 DNA	
<220> <221> <223>	primer_bind	the PSMA genomic the corresponding
<400> agaatggggt	29 ttagtttaat gg	22
<210> <211> <212> <213>	21	
<220> <221> <223>	primer_bind	the corresponding
<400> tgagtcactt	30 tttggagtca g	21
<210> <211> <212> <213>	22 DNA	
<220> <221> <223>	<pre>primer_bind sense oligonucleotide intronic sequences of clone used to amplify</pre>	the PSMA genomic
<400> ttgtaagcta	31 tccctataag ag	22
<210> <211>		

<212> <213>		quence	
<220> <221> <223>	primer_bind antisense oli intronic sequ clone used to	gonucleotide primences of the PSMA amplify the correspondence PSMA-like gene	genomic responding
<400> agttcagcaa	32 cagtcatgtt ag		22
<210> <211> <212> <213>	22 DNA	quence	
<220> <221> <223>	primer_bind sense oligonu intronic sequ clone used to	cleotide primer bences of the PSMA amplify the corr e PSMA-like gene	genomic esponding
<400> gggtggtcct	33 gaaaccaatc cc		22
<210> <211> <212> <213>	21 DNA	quence	
<220> <221> <223>	intronic sequ clone used to	gonucleotide primences of the PSMA amplify the correr PSMA-like gene	genomic esponding
<400> gtgatattac	34 agaaaggagt c		21
<210> <211> <212> <213>	22 DNA	quence	
<220>			

<221> <223>	sense oligonucle intronic sequenc clone used to am	otide primer based upon es of the PSMA genomic plify the corresponding SMA-like gene (exon 19)
<400> atccaggaat	35 tgcagagtgc tc	22
<210> <211> <212> <213>	22 DNA	nce
<220> <221> <223>	<pre>primer_bind antisense oligon intronic sequence clone used to am</pre>	ucleotide primer based upon es of the PSMA genomic plify the corresponding SMA-like gene (exon 19)
<400> ttcagtttta	36 atccataggg ag	22
<210> <211> <212> <213>	24 DNA	nce
<220> <221> <223>	primer_bind sense primer (ex	on 10) used for performing m various tissues
<400> acagatatgt	37 cattctggga ggtc	24
<210> <211> <212> <213>	DNA	nce
<220> <221> <223>	antisense primer	(exon 16) used for n cDNAs from various
<400> actgtgatac	38 agtggatagc cgct	24